Icesat Science Investigator-led Processing System (ISIPS) and Software Support

John DiMarzio, Raytheon ITSS
Anita Brenner, Raytheon ITSS
John Bay, Raytheon ITSS
Mark Sherman, Raytheon ITSS
David Hancock NASA/WFF Code 972
John.DiMarzio@gsfc.nasa.gov
301-614-5893

Science Data Centers Symposium: March 26-28, 2001
Pasadena CA

ICESat Spacecraft

- Single Instrument — Geoscience Laser Altimeter System (GLAS)
- Mission designed to measure
  - ice-sheet topography and associated temporal changes
  - cloud and atmospheric properties
  - land and water along-track topography
- Launch December 15, 2001

Raytheon ITSS Providing Broad Support of ICESat Mission

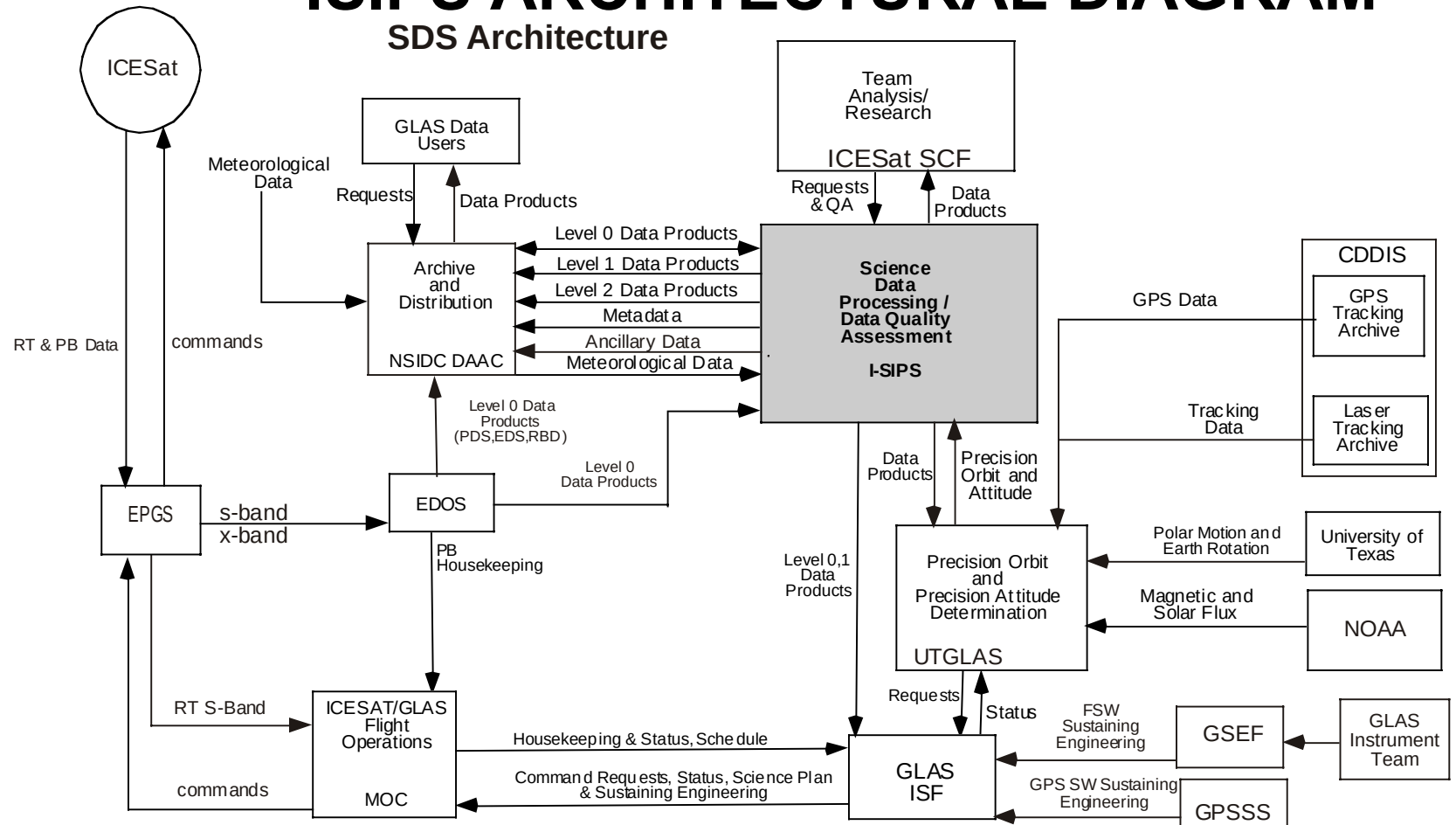
- Algorithm Support
  - Altimetry algorithm development
- Systems Engineering
  - Science Computing Facility design, implementation and maintenance
- Software Engineering
  - All ISIPS, SCF, and GLAS algorithm software
- GLAS Instrument Support
  - Running Instrument Support Terminal (IST).
  - Validating flight software.
- ISIPS Operations
- Data Management and Distribution to Science Team
- Data Visualization
- Science support for ICESat Project Scientist

GLAS SCF Software Components

- GLAS Science Computing Facility (SCF) Software
  - —Client-server data management, ordering, and visualization software
  - —Geographic and Temporal Subsetting
- GLAS Science Algorithm Software (GSAS)
  - —Creates GLAS standard products (controls which products are created and implements ATBDs)
- ICESat Science Investigator-led Processing System (ISIPS) Software
  - —Scheduling and Data Management Subsystem (SDMS)-Processing environment to control job flow, data distribution, and archiving

ISIPS ARCHITECTURAL DIAGRAM

SDS Architecture

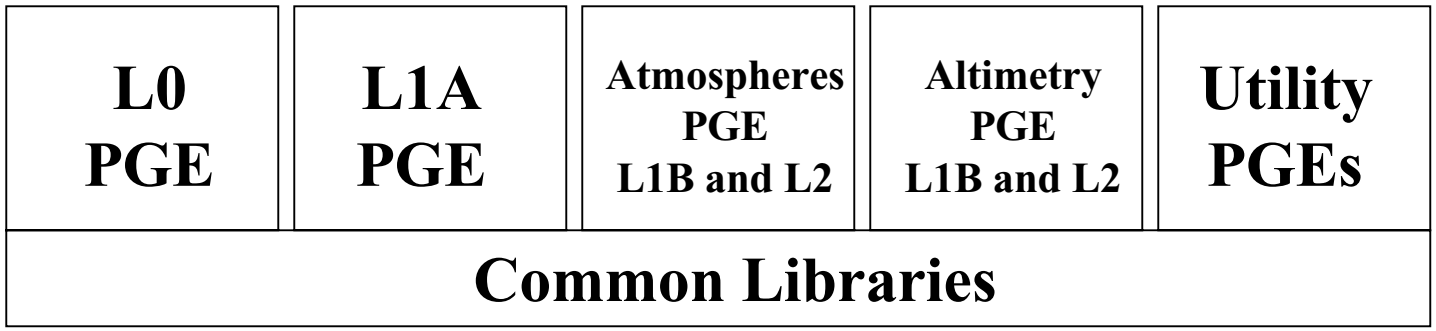


CDDIS	Crustal Dynamics Data and Information System
EDOS	EOS Data and Operations System
EOSDIS	EOS Data and Information System
EDS	Expedited Data Sets (X-band)
EPGS	EOS Polar Ground Station
FSW	Flight Software
ISF	Instrument Support Facility
GPS	Global Positioning System
GPSSS	GPS Sustaining Support
GSEF	GLAS Sustaining Engineering Facility
MOC	Mission Operations Center
NOAA	Nationjal Oceanic and Atmospheric Administration
PDS	Prodution Data Sets (X-band)
RBD	Rate Buffered Data (S-band)
SCF	Science Computing Facility

GLAS SCF Science Team Support

- Main SCF at GSFC will store all level 2 and selected level 1a & 1b products.
- Science Team members will view, subset and order data based on geographical and temporal regions of interest on local computers through the Main SCF
- All data products are available to science team members. Those not stored on SCF will be retrieved automatically from ISIPS archive.
- Data Selection, subsetting, ordering and distribution subsystem GUI
- Data Visualization subsystem

ISIPS GLAS Science Algorithm Software



- Produces all Level 1 and 2 GLAS standard data products
- PGEs produce interrelated product bundles in a single run
- PGEs run separately — can process concurrently
- Interface between PGEs and the SDMS utilize a keyword=value control file which defines all input/output for the run

High Level Requirements for ISIPS

- Create GLAS Level 1 and 2 Data Products
- Scheduling Requirements
  - 24x7 operation with minimal operator intervention
  - Automatically run PGEs when needed inputs available
  - Provide effective error recovery procedures
  - Allow ISIPS staff to monitor, control, prioritize workload
  - Manage system resources (Disk, CPU) for best throughput
- Data Management Requirements
  - Keep copy of all data products produced
  - Retain processing history for long-term problem analyses

Additional ISIPS Goals

- Scheduling
  - Integrated scheduling across multiple computer platforms (primary and backup)
- Data Management
  - Maintain two copies of all data products for additional protection against data loss for the life of the mission

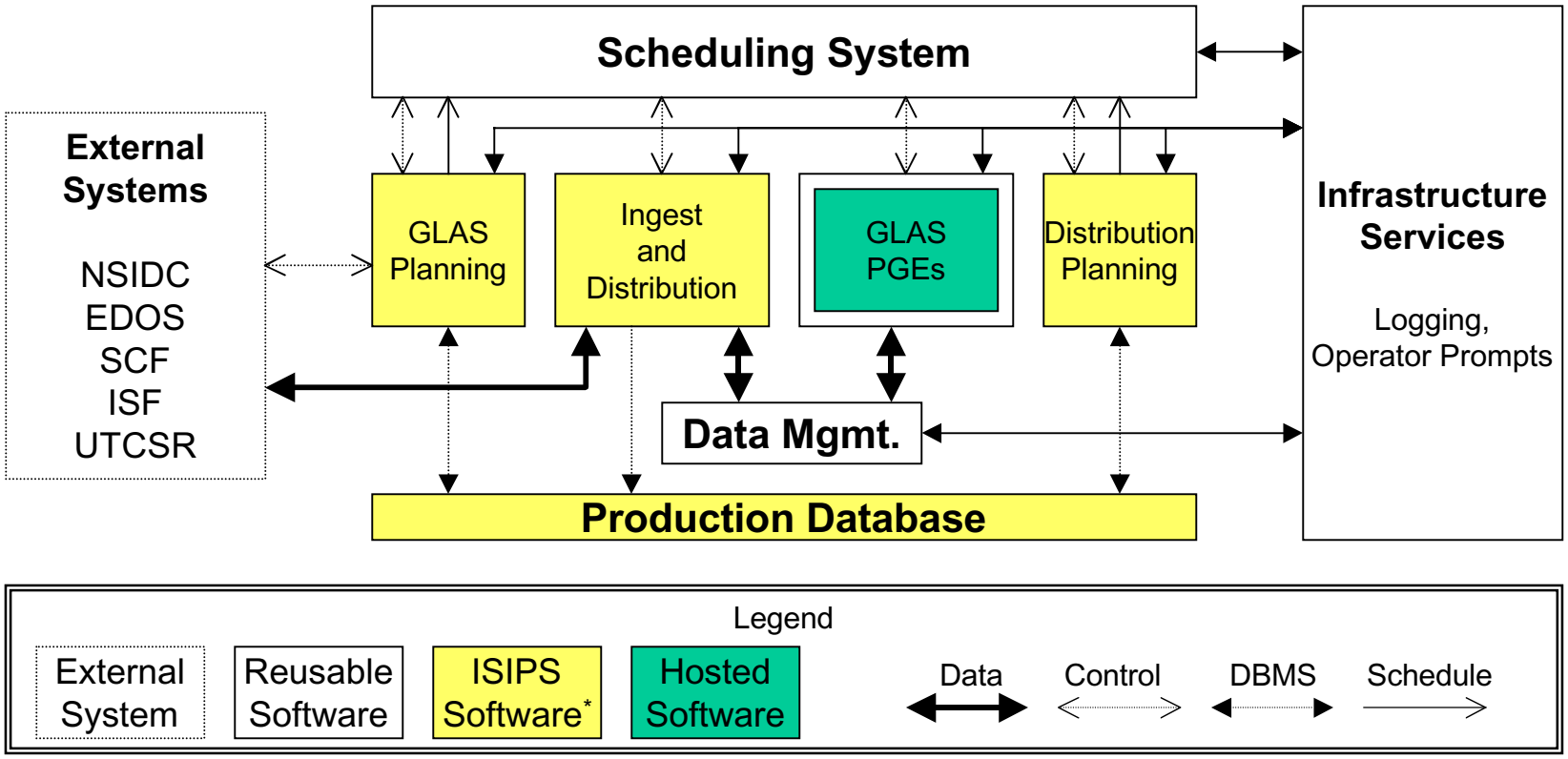
Scheduling and Data Management Subsystem (SDMS)

- The SDMS is software developed for ISIPS that:
  - Ingests data from external systems
  - Automatically runs PGEs on ingested data to produce GLAS products
  - Archives products into robotic digital libraries
  - Tracks all products ever created
  - Reprocesses data when new versions of science algorithms or new supporting data sets are released
  - Automatically distributes data to the NSIDC DAAC for further distribution to end users
  - Automatically distributes data to the Main GLAS SCF
  - Has high potential reuse in other science missions

SDMS Components

- The SDMS combines newly developed software with reused components of the GSFC V0 DAAC system.
  - Planning System
    - Written specifically for ISIPS using JAVA and JDBC with Oracle database
  - Scheduling System
    - Most elements ported and reused from V0 DAAC
  - Data Management System
    - New JAVA data server written using DAAC Distribution Cache Manager (DCM) as a model
    - Archer (tape library archive manager) ported to HP and reused
  - Distribution System
    - New external interfaces written
    - Distribution planning system written

SDMS Functional Architecture



ISIPS Planner: Job Initiation

The planner for a given bundle initiates a processing job when:

¥Each output product release for the bundle has unprocessed granules

¥Each input product release for the bundle has processed granules whose start and end times cover the outputs

When these conditions are met, the planner generates:

¥A job script to run the appropriate PGE

¥A control file containing required arguments (e.g. input and output file names) for the utility

The job script is submitted to the ISIPS Dispatcher, part of the Scheduling and Data Management Subsystem (SDMS).

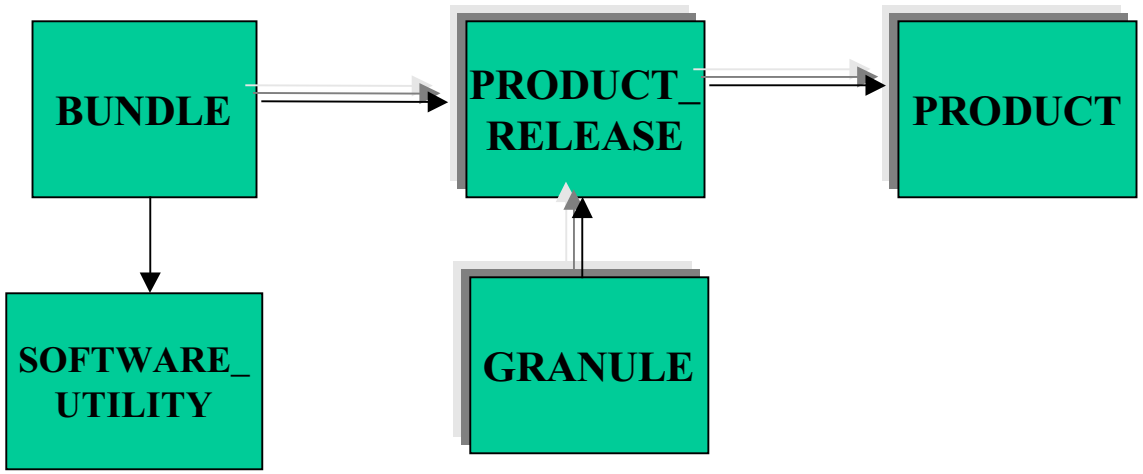
ISIPS Planner: Bundles

¥ **Each invocation of an ISIPS science utility produces granules belonging to a BUNDLE of products**

e.g. Elevation Processing produces granules for GLA06 and GLA12-GLA15

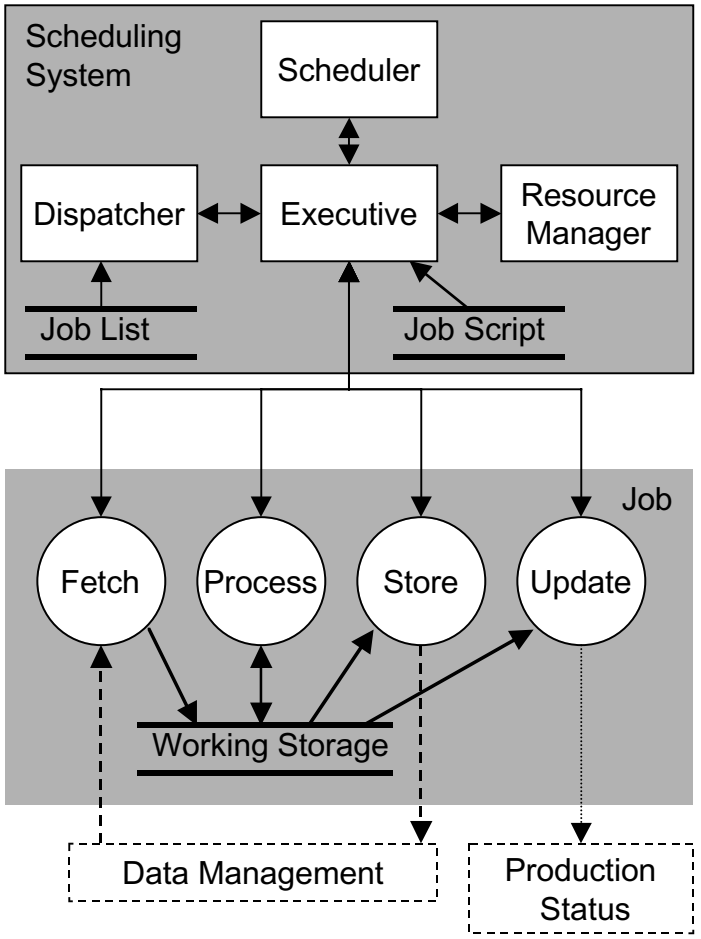
¥**The ISIPS Data Base records the input and output product releases associated with each bundle, and the granules associated with each product release.**

¥**Each bundle is also associated with the PGE that does the science processing.**



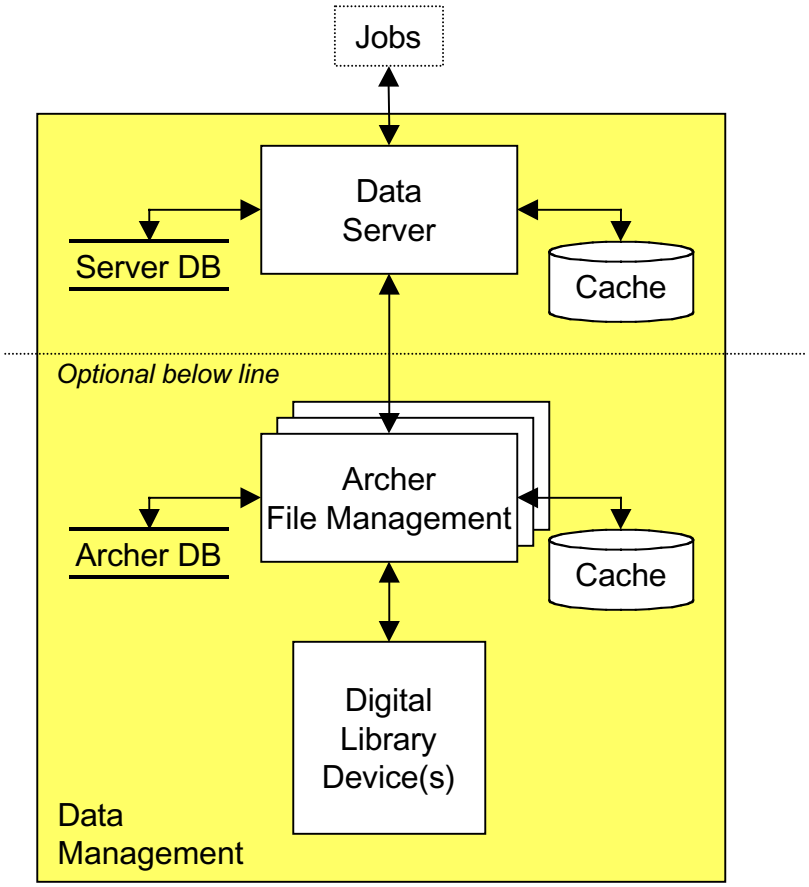
SDMS Scheduling System

- Scheduling runs jobs and steps
  - Steps are simply Unix processes
  - Jobs are steps that run together
  - Steps share a working directory
  - Scripted steps are in series or parallel
  - Multiple jobs can be run at once
- Scheduling has four components (most with GUI displays)
  - Dispatcher starts and tracks jobs
  - Executive controls active steps
  - Scheduler controls active jobs
  - Resource Manager allocates resources
- Typical processing job has four steps
  - Fetch to get data from Data Mgmt.
  - Process the data (science application)
  - Store results in Data Management
  - Update production status in DBMS

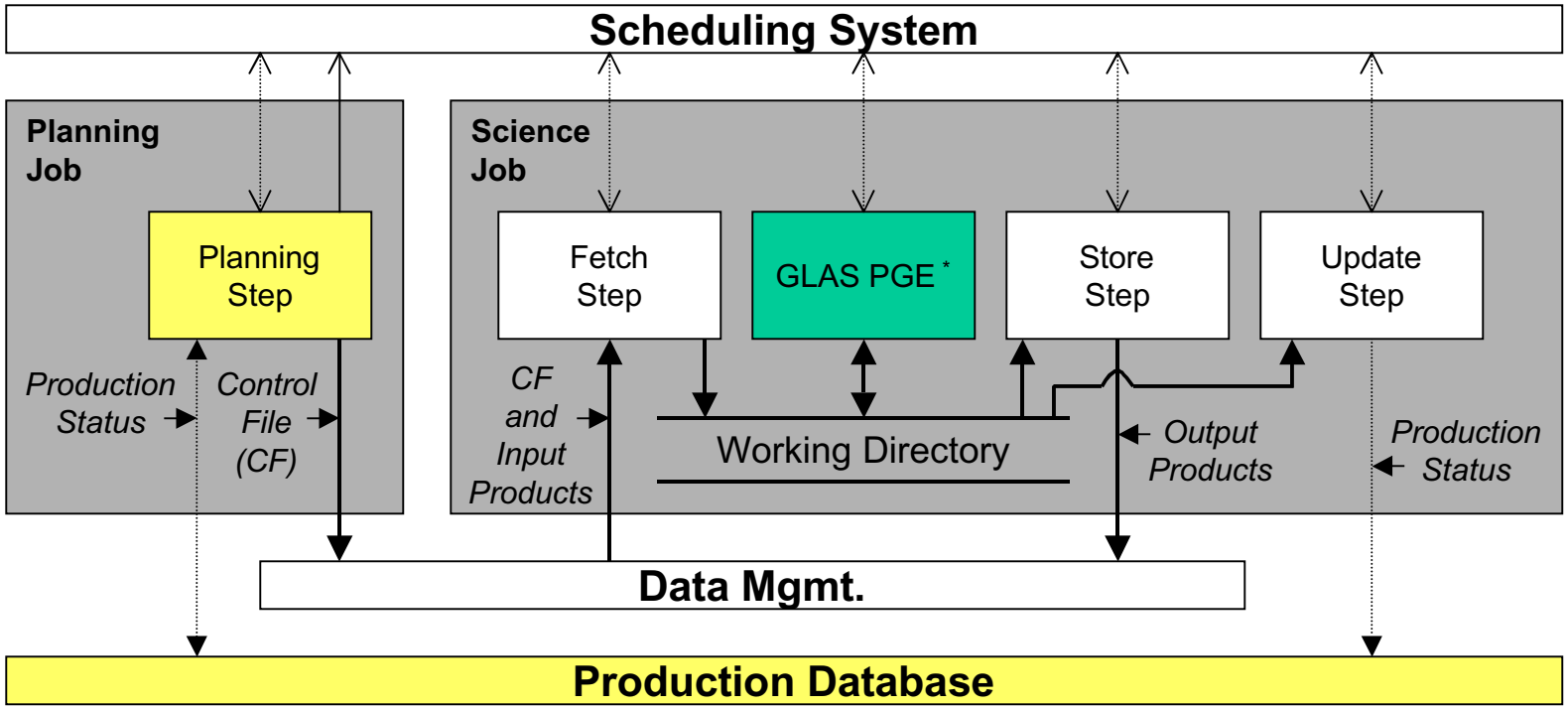


SDMS Data Management

- Data Management is a large repository for product files and other important data files
- Files can be held on disk, in near-line storage, or off-line
- Files are fetched/stored by name, location is transparent to client job
- Data server is high performance caching system
- ARCHER is optional archive file management system that moves files to/from digital library storage
- Use ARCHER system (or other HSM) only if more data is to be held than disk space is available
- Multiple archive systems like ARCHER may be used if desired



Science Processing Planning and Execution Model



* GLAS PGE is a Unix process, accepts command line from executive, does all I/O to working directory, returns 0 for success, non-zero for failure. Reads control file (as do the other steps) to determine what to do. These are the *only* interface requirements to run science processing in SDMS (i.e., no toolkit interface required).

Distribution Planning

- Driven by production database
  - —Reads through tables to find data ready for distribution
  - —Distributes data only if ruleset for the particular interface is satisfied
- External interfaces are being developed in JAVA for
  - —EDOS — incoming level 0 data
  - —GSFC DAAC — incoming meteorological data
  - —NSIDC DAAC — outgoing all level 1 and 2 data and metadata
  - —UT-CSR — incoming precision orbits and attitude files
  - —GLAS SCF — outgoing level 2 and selected level 1a and 1b data
  - —GLAS Instrument support terminal — outgoing GLAS housekeeping data

Conclusions

- We started with a reusable suite of software from the GSFC V0 DAAC
- We have produced an even more reusable suite of software for science processing
  - **Scheduling system can be reused**
  - **Data Management system can be reused**
  - **Infrastructure Services can be reused**
  - **Selected interface components can be reused (e.g., DAAC, EDOS)**
  - **Production Planning and Status needs to be developed on a per project basis (but model can be reused)**
  - **Science software plugs in very easily**